ETL-0469

DTIC FILE COPY

AD-A192 828

Linear Feature Extraction from Radar Imagery: SBIR Phase II, Base Contract Report

Gary D. Connor
Daryl T. Lawton
Christopher C. McConnell
David L. Milgram

Advanced Decision Systems
201 San Antonio Circle, Suite 286
Mountain View, CA 94040-1289

July 1987

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

Prepared for
U.S. ARMY CORPS OF ENGINEERS
ENGINEER TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VIRGINIA 22060-5546

88

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188
Exp Date Jun 30, 1986

1a REPORT SECURITY CLASSIFICATION
Unclassified

1b RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY
None

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE
None

3. DISTRIBUTION/AVAILABILITY OF REPORT
Distribution Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
TR-3139-01

5. MONITORING ORGANIZATION REPORT NUMBER(S)
ETL-0469

6a. NAME OF PERFORMING ORGANIZATION
Advanced Decision Systems

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
U.S. Army Engineer Topographic Laboratories

6c. ADDRESS *(City, State, and ZIP Code)*
201 San Antonio Circle
Suite 286
Mountain View, CA 94040-1289

7b. ADDRESS *(City, State, and ZIP Code)*
Code: CEETL-RI-A
Fort Belvoir, VA 22060-5546

8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Engineer Topographic Laboratories

8b. OFFICE SYMBOL *(If applicable)*
CEETL-RI-A

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
DACA72-86-C-0004

8c. ADDRESS *(City, State, and ZIP Code)*
Code: CEETL-RI-A
Fort Belvoir, VA 22060-5546

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO

11. TITLE *(Include Security Classification)*
Linear Feature Extraction from Radar Imagery, SBIR Phase II Base Contract (UNCLASSIFIED)

12. PERSONAL AUTHOR(S)
Gary D. Conner, Daryl T. Lawton, Christopher C. McConnell, David L. Milgram

13a. TYPE OF REPORT
Technical Report

13b. TIME COVERED
FROM 6/16/86 TO 1/15/87

14. DATE OF REPORT *(Year, Month, Day)*
July 1987

15. PAGE COUNT
39

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
SAR, Feature Extraction, Edge Detection, Terrain Analysis, Image Understanding

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

The goal of this effort is to develop and demonstrate prototype processing capabilities for a knowledge-based system to automatically extract and analyze linear features from synthetic aperture radar (SAR) imagery. This effort constitutes Phase II funding through the Defense Small Business Innovative Research (SBIR) Program. Previous work examined the feasibility of and technology issues involved in the development of an automated linear feature extraction system. The current effort continues this examination and is developing the technologies involved in automating this image understanding task. (Keywords)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
E. James Books

22b TELEPHONE *(Include Area Code)*
(202) 355-3039

22c OFFICE SYMBOL
CEETL-IM-T

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

PREFACE

This report describes work performed under contract DACA72-86-C-0004 for the U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia 22060-5546 by Advanced Decision Systems, Mountain View, California 94040-1289. The Contracting Officer's Technical Representative has been Dr. P. F. Chen.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Advanced Decision Systems (ADS) is pleased to submit this final technical report on research undertaken during the first part (Base contract) of a three part two year effort. The goal of this effort is to develop and demonstrate prototype processing capabilities for a knowledge-based system to automatically extract and analyze linear features from synthetic aperture radar (SAR) imagery. This effort constitutes Phase II funding through the Defense Small Business Innovative Research (SBIR) Program. The previous Phase I (contract DACA72-84-C-0014) work examined the feasibility of and technology issues involved in the development of an automated linear feature extraction system. The current Base contract effort continues this examination and is developing the technologies involved in automating this image understanding task.

2. EXECUTIVE SUMMARY

2.1 BACKGROUND OF PROBLEM

A vitally important problem facing the Department of Defense is the ability to quickly and efficiently analyze remotely sensed image data. This analysis is used for a variety of applications ranging from automated map making/updating to a variety of surveillance tasks, to other military and commercial remote sensing applications. An increasingly important and useful sensing capability is provided by synthetic aperture radar (SAR) imagery.

Imaging radar sensors provide all-weather, cloud penetration capability for a variety of applications. Technical capabilities now allow enormous volumes of such imagery to be automatically produced in relatively short periods of time. However, the current methods for analysis and interpretation of radar imagery largely consist of manual examination by human experts. As the quantity of imagery expands, the requirements for timely and efficient feature classification and the scarcity of radar image interpreters point to the need for an automated system for feature extraction and classification.

Linear features such as roads, rivers, bridges, and railroads are major landmarks in such imagery. Extracting and analyzing such features are prerequisites for most analysis applications. Traditional linear feature extraction techniques (edge detection and region segmentation) tend to perform adequately for low noise, high resolution visible imagery. However, the relatively poor quality and the complexity of the observed scenes in radar imagery make these feature extraction techniques less effective.

The ability to automatically detect and analyze linear features will have a major payoff for numerous applications. Technology to provide such an automated capability is emerging from the fields of image understanding (IU) and artificial intelligence (AI). Such a system could incorporate knowledge about the scene and use context (from the image or external sources such as digital terrain maps or terrain object models) to intelligently guide and interpret the extraction process. The results of the Phase I effort were encouraging in showing the feasibility of this approach. An automated system would greatly enhance the Army's capability for aerial cartography, change detection, aerial surveillance, and autonomous navigation. The goal of this effort is to pave the way for such a system by developing a largely automated terrain/image analysis workstation prototype.

There has been much work in artificial intelligence, computer vision, and graphics that satisfy the individual requirements for object modeling capabilities. Little has been done to integrate them, especially for the domain of SAR imaging. To date, the only vision systems that can interpret natural scenes deal only with very restricted environments [Hanson et. al. - 78] while other systems are restricted to artificial objects and environments. A system which used well defined shape attribute inheritance between a set of progressively more complex object models, and affixment relations that could be generalized to handle uncertainty begins to fulfill the basic requirements. It must also generate constraints on image features from object models. Care must be taken so that constraints on image structures generated from the abstract instances of object models are specific enough to generate initial correspondences between models and image structures. A rich set of image feature descriptions and robust object models that can adjust the segmentation process directly during their instantiation are also crucial to an automated system. Object models will be produced by ADS during the Option II phase of this effort for a limited set of features. A minimal object model must be able to direct constrained searches against image data. Models must eventually be capable of supporting learning and handling uncertainty in the matching of image feature descriptions to multiple terrain features.

The basic motivations for such a system stem from the poor results associated with the undirected application of low level image processing techniques. Environmental objects such as roads and rivers are semantic entities whose extraction requires contextual and object-specific knowledge which cannot be easily incorporated into, for example, low level filtering operations. In fact, it has become clear that a general and expandable system will have to incorporate processing which reflects the actual reasoning involved in expert SAR image interpretation.

The purpose of the Base Contract effort has been to undertake and complete the design of an automated linear feature extraction system for SAR imagery. The work performed by ADS in pursuit of this goal falls within three tightly coupled areas.

The primary work area focused on the continuation of the design produced in the Phase I SBIR effort. The results of that design are described in this document.

The second major work area was the design and development of a software environment within which to perform experiments and to build the eventual prototype system. The basic framework of this software was delivered to ETL in May. The delivery provided fundamental neighborhood and display operations. The software also contained the necessary software "hooks" for future expansion into the other system components.

Finally, the last concentration area was the continued experimentation with the government provided radar imagery. Experimentation included algorithm surveys, hand processing of sample imagery, and algorithm implementation. This work along with ADS's general understanding of machine vision, supports the design and development of the components of a model based vision system for linear feature extraction.

2.2 APPROACH

The major steps of this effort are as follows:

1. Develop the appropriate working environment to register, manipulate, and process imagery
2. Develop and experiment with various segmentation and feature extraction algorithms
3. Determine significant terrain object feature properties and construct representative object models
4. Experiment and evaluate model to image feature matching schemes
5. Develop an approach for managing competing and conflicting hypothesis matches
6. Develop feature finders/predictors to support or contradict an expected terrain feature's existence.
7. Implement a display interface to support the above processing steps.

Once the proper environment is established, the system for determining and extracting terrain features can be extensively tested. These experiments will further establish the role of autonomous feature extraction from SAR imagery and, indeed, the importance of SAR imagery to map generation.

2.3 PROGRESS TO DATE

2.3.1 Phase I

The major accomplishments of the Phase I effort were:

- Reviewed and implemented several edge and region extraction routines from optical image processing on SAR aerial imagery. Routines were evaluated for their performance in order to determine which would be valuable for integration into the general system.

- Obtained a better understanding of the nature of SAR aerial imagery and its requirements for interpretation.

- Considered a variety of techniques for representing the properties of environmental objects such as roads and rivers in SAR imagery.

- Designed and began component implementation of a model-based vision system for the extraction of linear features from SAR aerial imagery. In particular, ADS implemented an initial image structure data base and experimented with associated perceptual grouping rules and simple SAR object models.

A comprehensive report of Phase I results is available [Lawton et. al. - 85].

2.3.2 Phase II - Base Contract

The work performed by ADS under the Base Contract addresses three different problem areas.

The primary work area focused on the continuation of the design produced in the Phase I SBIR effort. The results of that design are described in Sections 3, 4, and 5 of this document.

The second major area in which ADS pursued the project goals was the development and the design of a software environment in which to perform experiments and begin to build the eventual prototype system. The basic framework of this software was delivered to ETL in May. The delivery emphasized

neighborhood and display operations. The software also contained the necessary software "hooks" for future expansion into the other system components.

Finally, the last area of work undertaken as part of the Base Contract was the continued experimentation with the government provided radar imagery. Experimentation included algorithm surveys, hand processing sample imagery, and actual algorithm implementation. This work and ADS's general understanding of machine vision, has been continually supporting the design and development of the components of a model based vision system for linear feature extraction.

2.4 ORGANIZATION OF THIS DOCUMENT

Section 3 recaps the research goals and briefly describes some of the major challenges of building an automated vision system.

Section 4 contains an overview of the system architecture which includes discussion on the system databases and the various inference processes. Three of the major components of this architecture are then highlighted in the following sections. Section 5.1 describes in detail the Perceptual Structure Database (PSDB - formerly the Image Structure Database, ISDB). Section 5.2 focuses on the description of Schemas, which are the hypotheses about world objects. Section 5.3 touches on the advantages of having underlying terrain maps to assist in the image exploitation process.

Section 6 closes with the status of the Phase II effort and a recap of the accomplishments of the Base Contract effort.

3. APPROACH

3.1 GOALS

The automatic extraction of linear terrain features from SAR imagery is the object of this research. Skilled analysts can perform this kind of task in the presence of very little context, e.g., no prior or ancillary imagery, maps or collateral descriptions. The analyst responds to simple visual cues in the imagery and forms hypotheses about their explanations, confirming or discarding them as the evidence requires. Analysts can fill in missing detail based on local cues and global hypotheses. The result of this effort is a map-like description of scene content, including road networks, rivers, bridges, field boundaries, etc.

Approaching human performance with an automated system requires an extensive knowledge base of expertise in detecting and recognizing instances of visual cues and a sophisticated inference mechanism Accordingly, the research in this project will attempt to:

- Develop perceptually-based object and terrain models.
- Develop methodologies to predict scene and image content from object and relationship models.
- Quantify perceptual laws of image feature aggregation for recognition of objects.
- Develop techniques to deal with uncertainties in prediction and recognition processes.
- Develop effective architectures for model-based vision.
- Integrate predictions with planning, modeling and image processing techniques to achieve autonomous formation of image segmentation strategies.

3.2 ISSUES

The above research objectives are subject to a number of important criteria for terrain and object modeling capabilities. The following requirements will heavily influence the design of the overall system.

Descriptive Adequacy: The modeling technique should be capable of adequately describing the terrain features. This includes representing naturally occurring features as well as man-made objects. It should be a consistent representation that supports modular system development and uniform inference procedures that can operate over different types of objects at different levels of detail.

Recognition Adequacy: Terrain models should be manipulable for determining the SAR appearances of world objects and for controlling recognition processing. This involves the formation of general predictions of sensor derived features from the terrain model. Such predictions will often be uncertain and qualitative due to outdated or incomplete prior knowledge of the terrain.

Handling Uncertainty: An automated system of this type must be able to handle uncertainty from a wide variety of sources. Beginning with the incoming registration parameters that should accompany each image to the "final" matching of image features to world objects, uncertainty needs to be properly managed so that consistent statements about the image scene can be made.

Primitive Learning: As the system begins to exploit an image of a scene, it should be possible for the system to adjust or "calibrate" model parameters based on the current scene content. Calibration is useful because it extends the system model to the actual image characteristics and features of the current data set.

Fusion of Information: As newer information is obtained by the system it must be combined or "fused" with the a priori map information and previous image collections. Therefore, the process of information fusion has both static and temporal characteristics.

3.3 TECHNICAL APPROACH

This Phase II effort will complete the development of the system design from Phase I, incrementally prototype it, and demonstrate its use on an expanded set of environmental linear features. The core components of this system will serve as a common testbed for research at both ADS and ETL. Our research will develop a sequence of increasingly sophisticated linear feature recognition techniques based upon successively more general object representations and inference procedures. This sequence will provide a basis for incrementally prototyping the total system at different stages of its development.

More specifically, the project will satisfy its objectives through a combination of the following types of activities:

- Infrastructure Software Design and Development
- Application Software Design and Development
- Experiments with SAR Imagery and Applicable Context (e.g., Platform Paramenters and Terrain Databases)

Although infrastructure development will dominate the first half of the project lifecycle, there will be preliminary efforts ongoing in the other two areas as well. As the project matures, focus will shift to the design and execution of increasingly more competent experiments to extract linear features from the SAR image test set. The software developed within these experiments will form the core of the application system prototype. The overall design of the prototype will follow the structure described below (Section 4).

4. SYSTEM ARCHITECTURE

The system architecture consists of several databases and inference processes. The inference processes transform the databases, creating additional data structures, and modifying the existing ones. The task interface focuses attention in system processing and monitors progress toward system task goals. This high level architecture is depicted in Figure 4-1. The boxes with square corners in this figure represent databases, the ellipses represent inference processes, and arrows indicate dataflow. The remainder of this section provides additional detail on the various databases and inference processes.

4.1 SYSTEM DATABASES

At the highest level there are three databases. These are the short term memory (STM), long term memory (LTM), and generic models.

The STM acts as a dynamic scratchpad for the vision system. It has two sub-areas, a perceptual structures database (PSDB) and a hypothesis space. The PSDB includes incoming imagery from sensors, immediate results of extracting image structures such as curves, regions and surfaces, spatial/temporal groupings of these structures.

The hypotheses space contains statements about objects and terrain in the world. A hypothesis is represented as an instantiated schema. The schema points to the various perceptual structures in the PSDB that provide evidence that the object represented by the schema (such as a terrain patch, road, forest, etc.) exists in the world. A hypothesis with no associated perceptual structures is a prediction. As structures and localization are incrementally added to a hypothesis, it progresses on the continuum from predicted to recognized. Hypotheses that have enough evidence associated with them to be considered recognized and stable, are moved to the LTM.

The LTM stores a priori terrain representations, the long term terrain database, and hypotheses with enough associated evidence to be considered visually stable. A priori data concerning terrain type information, elevation, and knowledge of specific landmarks are stored in the LTM. Consistency of one hypothesis with another is not required for storage in the LTM, although it is a goal. The area of unresolved conflicts will be further investigated as the effort continues.

The model space stores generic object models, the inheritance relations of the (model) schema network, and a set of image structure grouping processes and

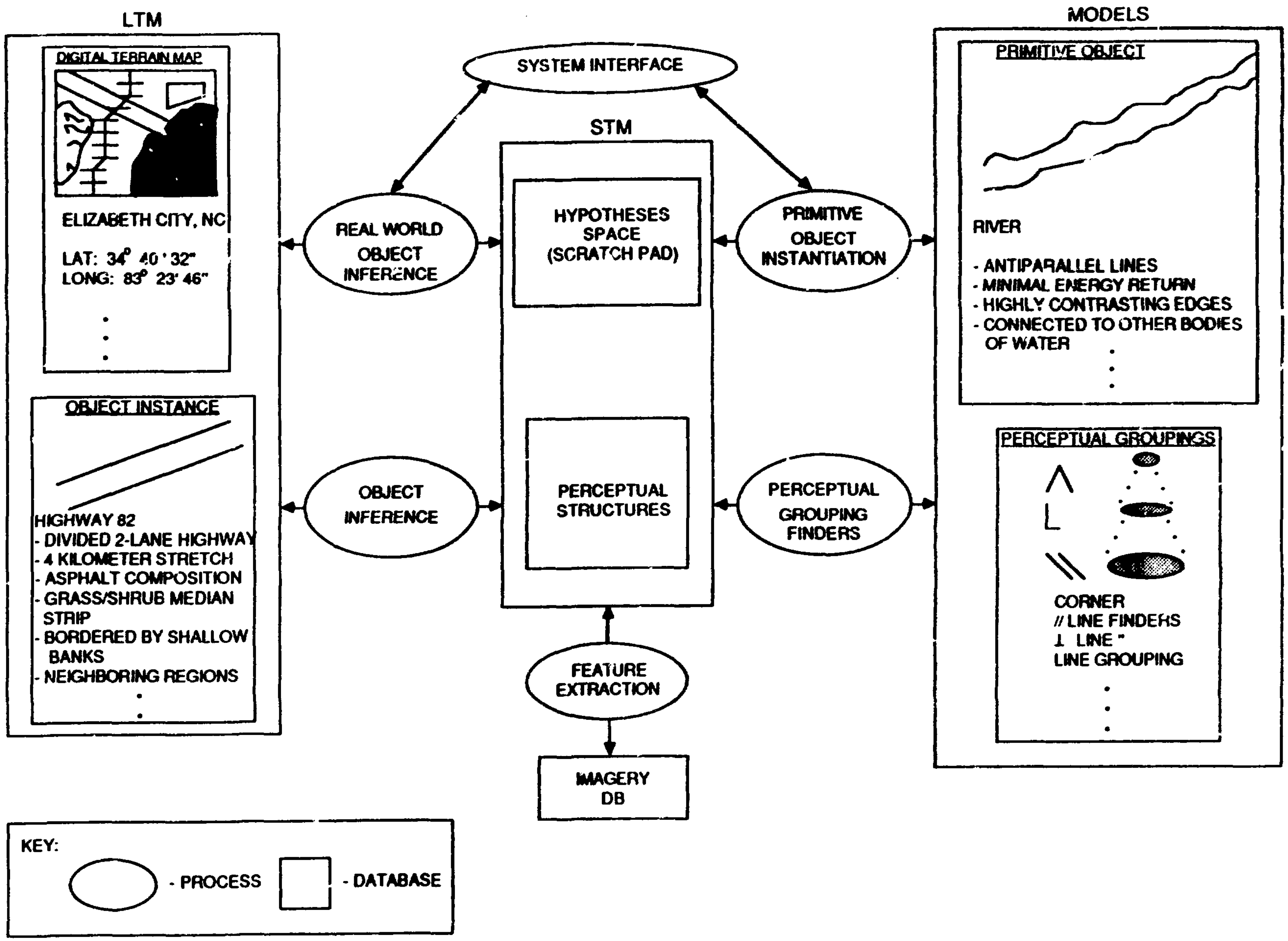


Figure 4-1: System Architecture

rules for evaluating image structure interestingness. Generic models are used dynamically to instantiate and guide search processes to associate evidence to an object instance. Inheritance relations are used by various schema inference procedures to propagate structures, attributes and relations between object instantiations. For instance, the generic two-lane-road schema has an "IS-A" relationship to the generic road schema. It follows, based on the inheritance models, that an instantiation of the two-lane-road schema will inherit the more general characteristics of the generic road schema that in turn inherits the more general characteristics of a terrain patch.

4.2 INFERENCE PROCESSES

At the highest level, there are five different sorts of inference processes in the vision system. These are perceptual inference, location inference, object instantiation, LTM/STM instantiation, and the task interface.

The PSDB is typically initialized with the output of standard image processing operations for smoothing, edge extraction, etc. Much subtler inference is required for grouping processes that produce connected curves, textures, and surfaces. These grouping operations are typically model guided. There are generic models (which may be task dependent) of what constitutes "interestingness" of an image structure.

The hypothesis inference processes produce tasks for the perceptual processes. These may be satisfied by simple queries over the PSDB such as "find all long lines in this region of the image", where "long", "line" and "region" are suitably interpreted. Queries can be arbitrarily complex, containing qualitative descriptors that are rigorously defined. Alternatively, the requested perceptual structures may be dynamically extracted. In this case, a history of the processing attempts and results are maintained. If similar requests are made later, such as if the system were to view the same environment from a different perspective, these processing histories could be used to recall a processing sequence that had produced successful results.

Location inference is basically the registration process. In downward looking optical imagery, the location of features in the scene is determined through the registration parameters. Generally, this is a simple determinate process (although, establishing the registration is not necessarily simple). However, accurate feature location in SAR imagery is somewhat problematic since the typical SAR viewpoint is side-looking at a narrow angle. There are therefore significant effects which can distort inferred locations. Among these are shadowing, layover, and complex scattering. However, if these effects can be detected, they reveal information about the local 3-D structure of the scene. Three dimensional information will be limited to a small set of features that are indicative of relative height; for

example, patches of dense forest cast measurable shadows that are indicative of a local change in height.

Generic schemas are models of world objects that include information and procedures on how to predict and match the object models to the available sensor data. Such schemas represent geometric constraints and qualitative sensor view appearance (image feature characteristics) including effects of change in resolution and environmental effects such as season, weather, etc.. Furthermore, schemas also indicate contextual relationships with other objects, type and spatial constraints, similarity and conflict relations, and spatial localization.

Object schema instantiation may occur by model-driven prediction from a priori knowledge, or directly from another instantiation and a PART-OF relation. The other instantiation process may also occur by matching a distinctive perceptual structure to a schema appearance instance. This sort of "triggering" is more common in situations where there is little a priori information to guide prediction, such as a lack of the underlying terrain map. Object instantiations generate queries to the PSDB grouping/searching processes in order to complete matching.

A key idea in object instantiation processing is inference over the model schema network hierarchies. Direct representation and inference over a large enough body of world objects to accomplish outdoor terrain understanding requires very large memory and proportionately lengthy inference procedures over that memory space. Hierarchical representation makes a significant reduction in storage requirements; furthermore, it lends itself naturally to matching schema to world objects at multiple levels of abstraction, thus speeding the inference process. Two basic hierarchies are the IS-A and PART-OF trees.

IS-A hierarchies represent the refinement of object classification. By having a hierarchy of classifications, image structures can be matched at multiple levels. The level to which an imaged object is classified is dependent upon the following:

- image information content (resolution, imaging angle, range, etc)
- segmentation ability of the system
- model descriptiveness (accuracy of the representation)
- matching of model descriptions to image features.

The benefit of the IS-A hierarchies stems from the capability to match the "resolution" of the information (see above) to the "resolution" of the appropriate level in the IS-A hierarchy. For example, an image may only be of sufficient resolution

to distinguish that a particular image feature is a road segment. Additional imagery may later lead the system to conclude that the previous "road segment" is actually a primary (highway or autobahn) class road. Even though the original classification of the image feature was "road segment" and it was later reclassified as a "primary road", the original classification was not wrong, just not fully classified. Figure 4-2 shows part of an IS-A hierarchy for terrain representation. All objects are world objects, this is the top node in the IS-A tree. The next level down in the tree represents a division into "MAN-MADE" and "NATURAL" objects. Continuing down the tree further classifies "MAN-MADE" objects as "RAILROADS", "ROADS", and other objects. Continuing down the "ROADS" branch reveils various types of roads. At the "MAN-MADE" level the IS-A tree provides only gross level information. Man-made objects typically have more structure than naturally occurring objects. The next level in the tree starts to provide additional descriptive information. Roads are typically anti-parallel sets of lines that have certain directional and spatial properties. Further traversal of the tree begins to reveal even more specific information. Information such as composition, number of lanes (size), expected number of branches, etc is inherited from the specific instances of the various road schemas.

Whereas IS-A hierarchies describe the world in ever finer levels of classification, PART-OF hierarchies represent the decomposition of world objects into components, each of which is itself another world object. PART-OF hierarchies contain relative geometric information that is useful in prediction and search. Figure 4-3 shows a PART-OF hierarchy decomposition for a generic "PRIMARY-ROAD".

As object instantiation inference reasons up and down schema network hierarchies, incrementally matching perceptual structures and other data to instances of object appearance in the world, a history mechanism records the inference processing steps, parameters and results. This dynamic data structure is called the schema instantiation structure. One important aspect of this structure is that it can be used to extract the inference and processing sequence(s) that worked earlier to see the same object, or ones that are similar. This accounts for the fact that distinctiveness in image appearance is dependent on many factors which are difficult to predict given a real world environment which contains many unknowns.

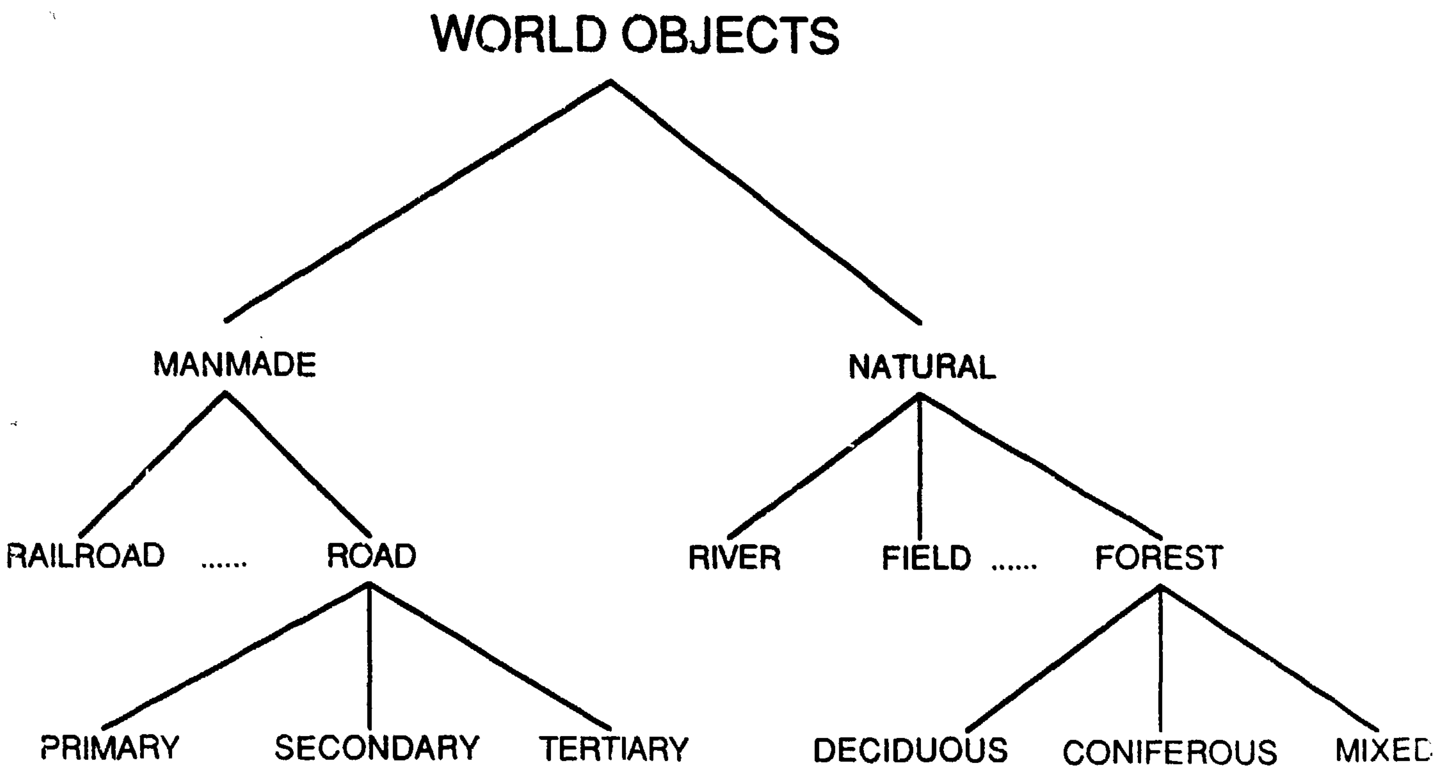


Figure 4-2: IS-A Hierarchy

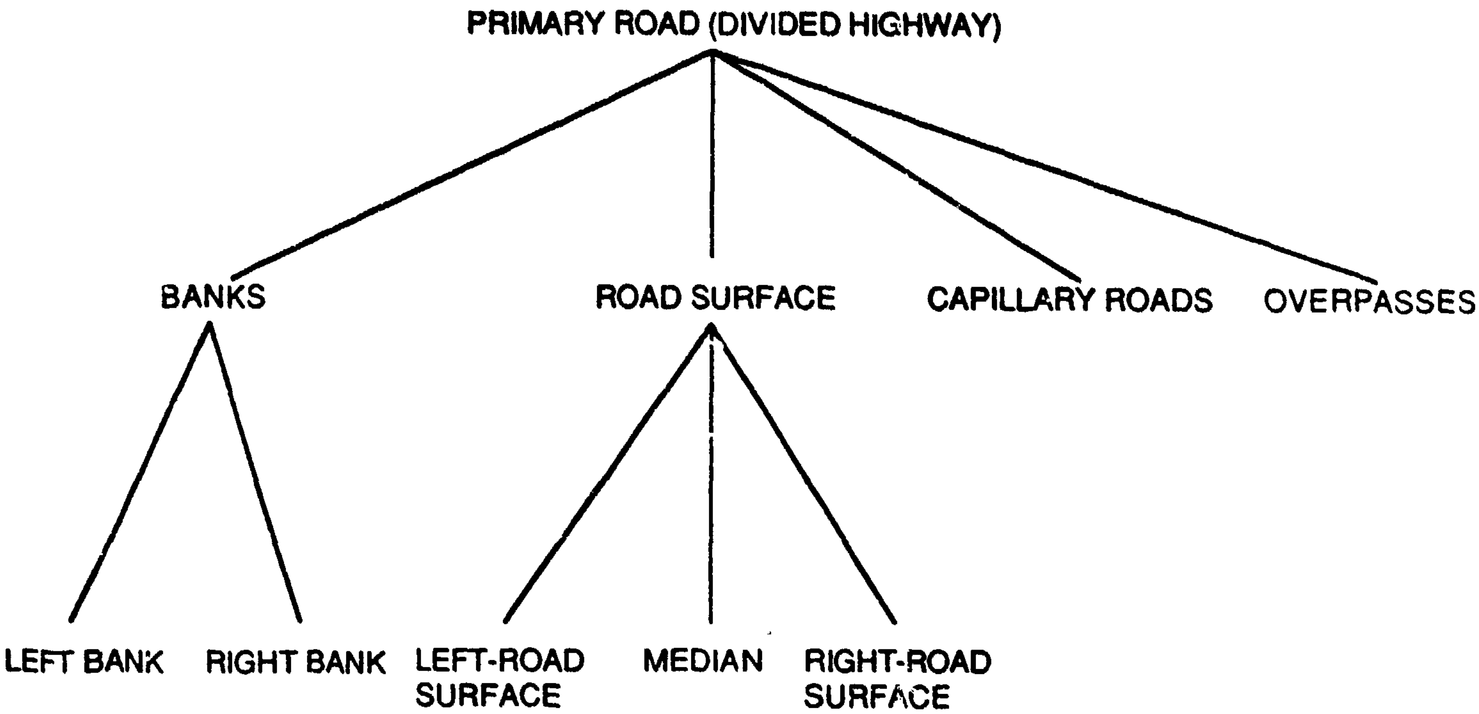


Figure 4-3: PART-OF Hierarchy

5. COMPONENT TECHNOLOGY

5.1 PERCEPTUAL STRUCTURES MANAGEMENT

Perceptual processing is concerned with organizing images into meaningful chunks. From a data-driven perspective, the definition of "meaningful" and the development of explicit criteria to evaluate segmentation techniques requires the chunks to have characterizing properties, such as regularity, connectedness, and fragmentation resistance. From a model-driven point of view, "meaningful" is defined as the extent to which chunks can be matched to structures and predictions derived from object models. From either perspective, a basic requirement is that image segmentation procedures find significant image structures, independent of world semantics, in order to initialize and cue model matching. This means that the extraction of image events such as surfaces, boundaries, and interesting patterns takes place independently of the context of a particular object. These extracted structures are, in turn, useful primitives both to match up with components of object models and to be used to describe the characteristics of novel (unmodeled) objects.

The Perceptual Structure Data Base (PSDB), conceptualized in Figure 5-1, contains several different types of information. These are classified as images, perceptual objects, and groups. Images are the arrays of numbers obtained from the different sensors and the results of low level image processing (such as smoothing operators or median filters) that produce such arrays. It is difficult for the symbolic/relational representations used for object models, such as schemas, and the processing rules in computer vision systems, to work directly with an array of numbers. Therefore, there are many spatially-tagged, symbolic representations used in image understanding systems that describe extracted image structures such as the primal sketch [Marr - 82], the RSV structure of the VISIONS system [Hanson et al. - 78], and the patchery data structure of Ohta [Ohta - 80]. Such a representation has been built around a set of basic perceptual objects corresponding to points, curves, regions, surfaces, and volumes.

Groupings are recursively defined to be a related set of such objects. The relation may be exactly determined, as in representing which edges are directly adjacent to a region, or they may require a grouping procedure to determine the set of objects that satisfy the relationship. Groupings can occur over space, e.g., linking texture elements under some shape criteria such as compactness and density, or over time, as in associating instances of perceptual structures in overlapping images.

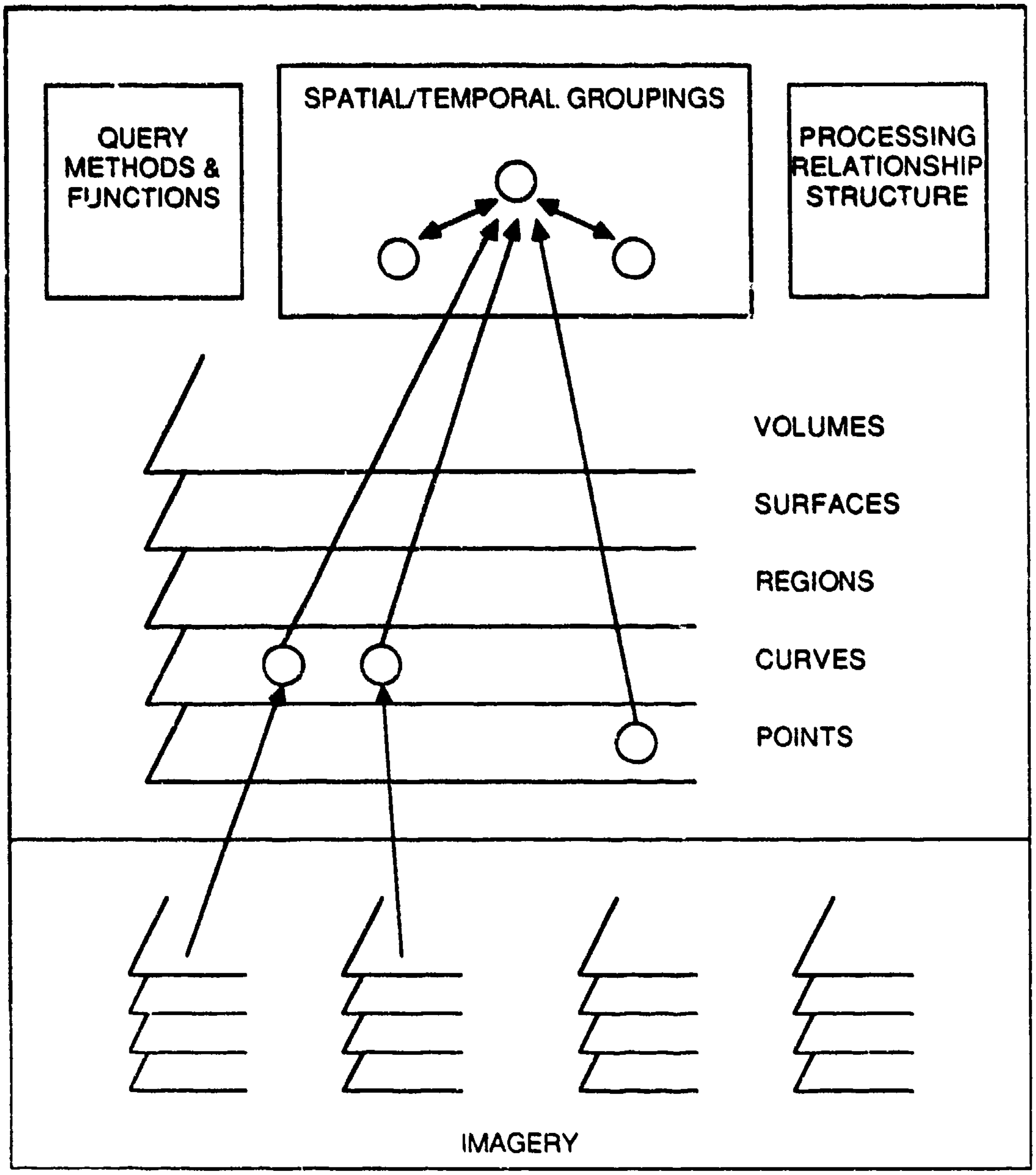


Figure 5-1: Perceptual Structure Data Base (PSDB)

5.1.1 Initialisation of the PSDB

Whenever new image data is obtained, a default set of operations are performed to initialize the PSDB. For example, edges could be extracted at multiple spatial frequencies and decomposed into linear subsegments. The edges would be extracted into distinct connected curves, and general attributes such as average intensity, contrast, and variance would then be associated with each. Similar processing could be performed for extracting regions. For example, histograms could be computed with respect to a wide range of object based and image based characteristics in a pyramid like structure. These default operations are used to initialise bottom-up grouping processes and schema instantiations. These, in turn, determine significant structures using heuristic interestingness rules to prioritise the structures for the application of grouping processes or object instantiations.

5.1.2 Images

Images are simply the data arrays derived from the imaging sensors, SAR in the context of this project. The results of image processing routines that produce arrays of data are also treated like images, processes like averaging, speckle reduction, and gradient computations. Associated with images are several attributes for time of acquisition, relevant sensor parameters, etc. Processing history is maintained in the processing relationship structure that keeps track of the processing history of all objects in the PSDB.

5.1.3 Perceptual Objects

Points, curves, regions, surfaces, and volumes are basic types of perceptual structures that are accessible to object instantiations and grouping processes. An example instance of a curve structure is shown in Figure 5-2. This figure shows many common representational characteristics of perceptual objects. There are required attributes associated with particular objects, such as endpoints, length and positions for a curve. There is also an associated attribute-list mechanism for incorporating more general properties with an object. This list is accessible by keywords and a general query mechanism using methods specific to the particular associated attribute. The associated attributes in the example are shown in capital letters. There are many types of attributes that can be consistently associated with a curve using this mechanism.

A useful representation for performing geometric operations and queries over objects is the OBJECT LABEL-GRID (or GRID: in the example curve. The number 6 indicates the index of this structure). This is an image where each pixel contains a vector of pointers back to the set of perceptual objects and groups

```
Curve: #<CURVE 175505263>
Point1: (205 291)
Point2: (274 285)
Length: 81
Grid: 6
Points: (205 291) (206 292) (207 293) ...... (276 287) (275 286) (274 285)
AVG-INTENSITY: 159.27315
SUCCESSIVE-GRADIENT-DIFFERENCE: 1.7660371 1.4782858 1.0026073 ......
                                0.5625856 0.6442404 0.7240415
GRADIENT: (-2.8555908 3.1515503) (-4.0343933 1.8365173) (-5.415619 1.3096924) ......
          (-0.03842163 -2.9117389) (0.56552124 -2.6931152) (1.0851135 -2.1888733)
CONTRAST: 5.7423205
AVG-GRADIENT: -5.7336416 0.31559697
LDC: (#<CURVE 175505520>
       (#<CURVE 175327156>)
       (#<CURVE 175505525>
          (#<CURVE 175327163>)
          (#<CURVE 175505532>
             (#<CURVE 175327170>)
             (#<CURVE 175327175>))))
```

Figure 5-2: Curve Example

which occupy that position. This allows geometric operations to be performed directly on the grid. Filtering operations can be applied to the OBJECT LABEL GRID to restrict processing based upon attributes associated with objects. Various types of masks can be associated with objects to reflect a directional or uniform neighborhood to determine object relationships in the OBJECT LABEL GRID.

5.1.4 Groups

A group is a set of related perceptual objects. The relation can be determined directly by a query over an object and those surrounding it, as in finding the set of curves within some distance of a given region. Alternatively, it may require a search process to find the set of objects meeting some, potentially complex, criteria. For example, an ordered set of curves can be grouped together using thresholds on allowable changes in the average contrast and orientation of successive elements. By expressing the grouping process as a search over a state space of potential groups, each group becomes a potential hypothesis in the PSDB. A relational grouping procedure is illustrated in Figure 5-3 for the determination of nearby parallel lines with opposite contrast directions. This is done for a linear segment by first extracting nearby neighbors using a narrow mask oriented perpendicular from the segment at its mid-point. The intersection of this mask with points in the label grid are determined, and then each candidate is evaluated by checking if it is within allowable thresholds for length, contrast, and orientation. It is then ordered with respect to the smallest magnitude of the difference vector computed from the average gradients. The grouping processes can either produce the best candidate as a potential grouping, or some set of them.

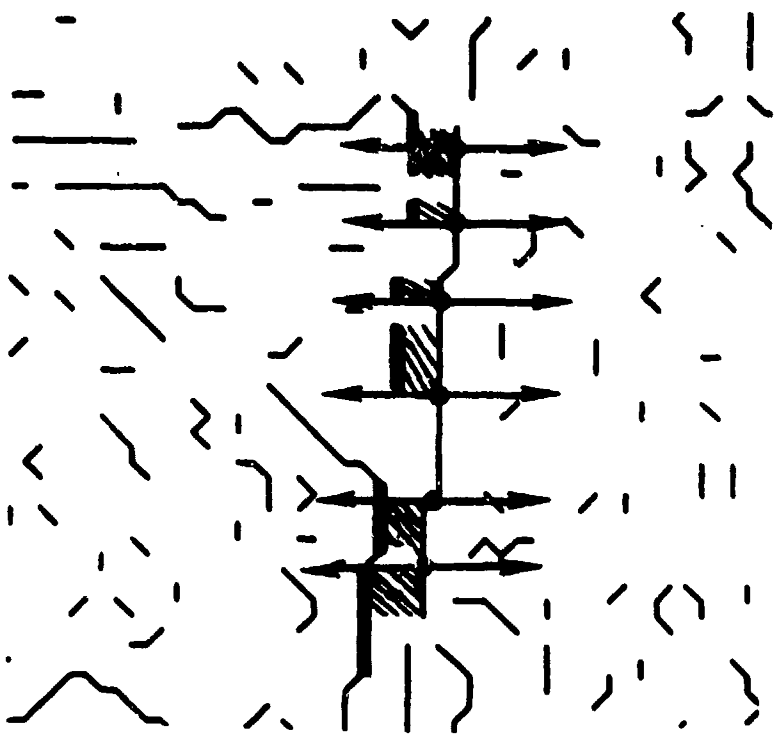
Two different types of grouping processes have been developed: measure-based and interestingness-based. The measure based grouper is a generalization of established edge and region linkers [Martelli - 76]. It uses a measure consisting of:

1) some value to be optimized, such as length, minimal curvature, compactness, or a composite scalar value

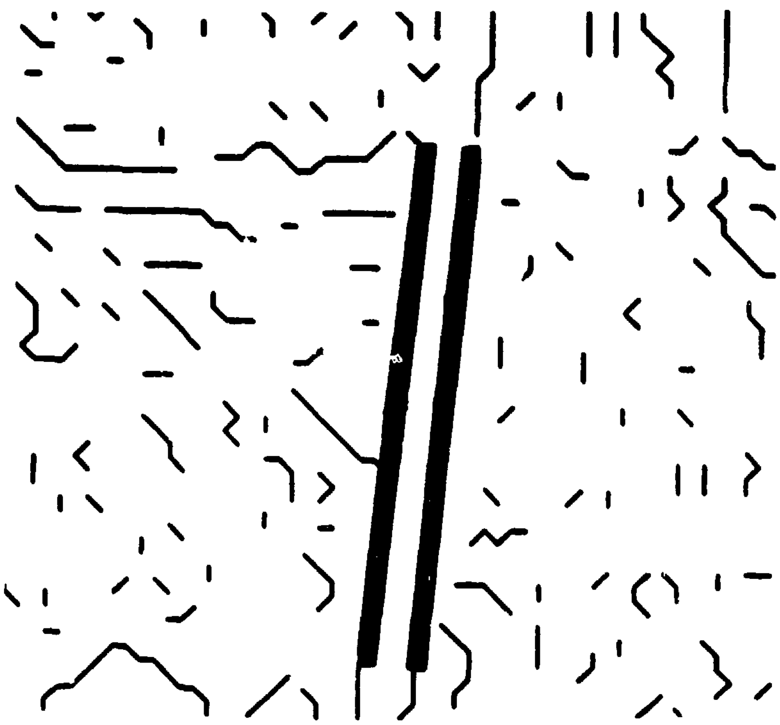
2) local constraints on allowable changes in attributes

3) global thresholds on attributes

The measure and associated constraints are optimized by a best first search returning several ordered candidate groups. The measure to be used can be



a) Edge-to-Edge Association



b) Result of Grouping Step

Figure 5-3: Parallel Grouping

associated with a prediction from an object model for substance or shape characteristics. The measure to be optimized can also be determined directly from initially extracted objects by selecting those that are extreme in some attribute or are correlated with the attributes of surrounding objects to derive a measure to be optimized.

The measure based grouper is currently being generalized into one based on interestingness. It involves the basic processing loop shown in Figure 5-4. Initially, basic perceptual objects including curves, regions, junctions and their associated attributes are extracted using conventional techniques. Extracted objects are represented in label grids to express spatial neighborhood operations over the objects. A uniform neighborhood is established for each object, and directed relations are formed with the adjacent objects in each neighborhood. These relations are represented in a small number of types of match relationships that contain descriptions of the correlation of attributes, subcomponent matching, and composite properties.

Selected attributes of the extracted perceptual objects and the match structures are then sorted into lists with pointers back to the associated objects. These lists are for attributes such as size, average feature values, variance of feature values, compactness, the extent of correlation between the components and attributes of different structures, and the number of groups an object is involved in. These different rankings are then combined using a selection criteria to choose the set of interesting perceptual objects and relationships. The selection criteria sets the required position in different subsets of the sorted attribute lists. An example is to find 100 largest objects in the top 10 of any of the attribute correlation lists. The selection criteria is modifiable during processing and is meant to reflect the influence of model-based predictions.

Interestingness is used to focus the application of grouping rules to a selected set of objects and relations between objects indicated in match structures. The grouping rules then combine perceptual objects to form new perceptual objects, or groups, based upon the type of relation between the objects. Neighborhoods are established with respect to these derived groups to form new relationships. These in turn are sorted in the attribute lists with respect to the previously extracted perceptual objects. In addition to the relations established in uniform neighborhoods, for some groups, non-uniform relations are also established. Processing can continue indefinitely as less and less interesting relations become candidates for the application of grouping rules. Explicit criteria are needed to stop processing; e.g., limiting processing time, determine when there is a uniform covering of the image with extracted groups, or when structures belong to unique groups.

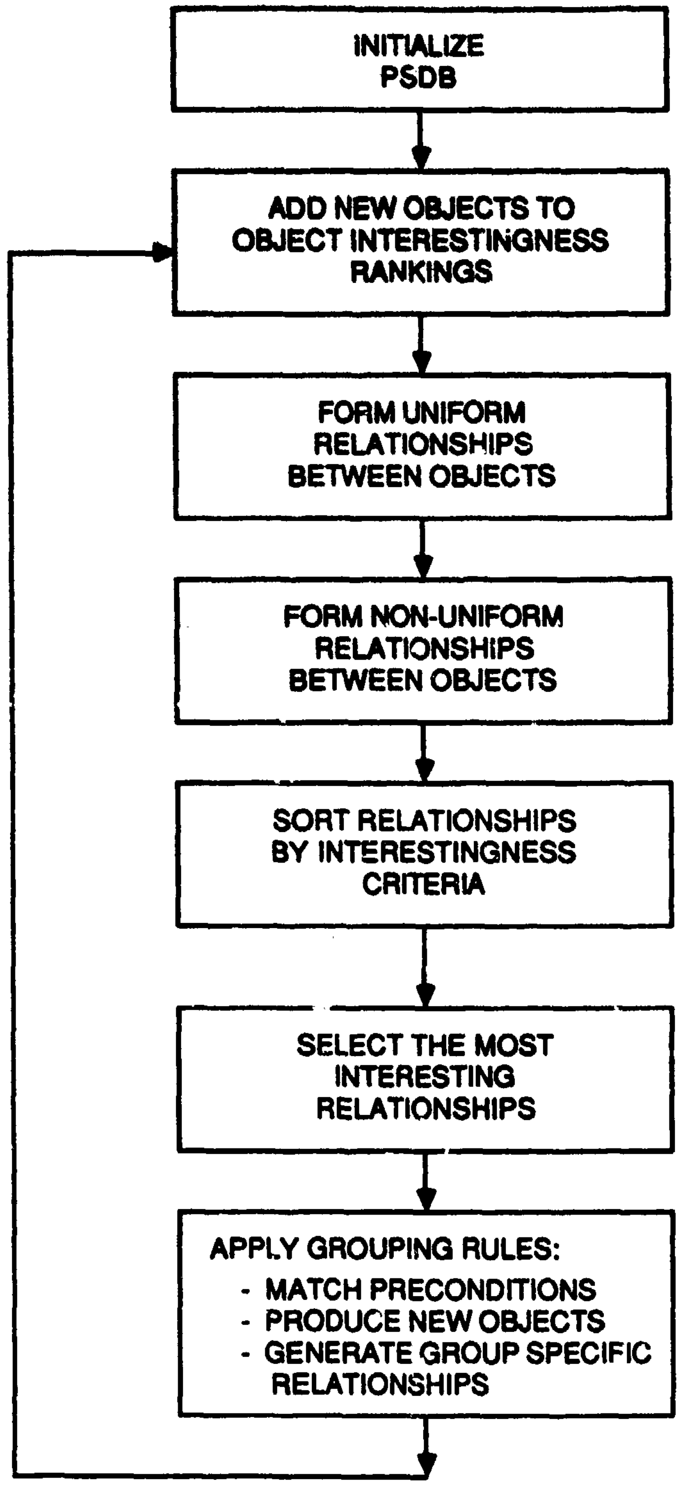


Figure 5-4: Grouping Processing Flow

These operations are performed by virtual processors called grouping nodes. Grouping nodes are seen as covering regular and adjacent portions of an image area The image area contains some portion of a label plane for accessing the objects based upon their spatial dispositions as well as object-based associated attributes. The grouping nodes are further organised in a hierarchical pyramid shown in Figure 5-5. Each node is connected to its adjacent neighbors and has a parent and descendants. The transfer of information between nodes at different levels is based upon interestingness. Lower level processes send their most interesting structures up the hierarchy. There are several effects of this. One is that it allows a uniform processing to occur at different levels, so grouping rules can be applied to objects at different levels of interestingness. It also allows relations between nonspatially adjacent structures to be handled in a uniform architecture. It also partitions perceptual structures in a way that corresponds to different levels of control in instantiation of object models.

Organising segmentation in terms of grouping processes has many advantages for a model based vision system. The grouping processes can be run automatically from extracted significant structures based upon perceptually significant, though non-semantic criteria. Thus, connected curves of slowly changing orientation or compact, homogeneous regions can be extracted purely on perceptual criteria. These image structures correspond to world structure and events, and they are useful for initialising schema instantiations. They correspond to the qualitative image predictions associated with more general schemas. An inference process for compilation from an object model into grouping processes, allows model based vision to have a very active character quite different from single-level attribute matching.

5.2 SCHEMAS AND RECOGNITION

Schemas represent hypotheses about objects in the world. A schema can represent perceived, but unrecognized, visual events, as well as recognized objects and their relationships in image scenes. The architectural design is focused about the representation, instantiation, and inference over schemas developed by the system. Schemas are related to similar concepts found in [Hanson et.al. - 78] and [Ohta - 80]. The hypothesis space found in short term memory (STM) consists of schema instantiations that represent accumulated perceptual evidence for objects as attributes and relations (labels) that are instantiated with varying levels of certainty.

Object models are used to organise perceptual processing by integrating descriptive representations with recognition and segmentation control. One aspect of this is the use of different types of attributes and inheritance relations between generic schemas for representation in IS-A and PART-OF hierarchies. These viewing attributes are also inherited and modified according to different object types.

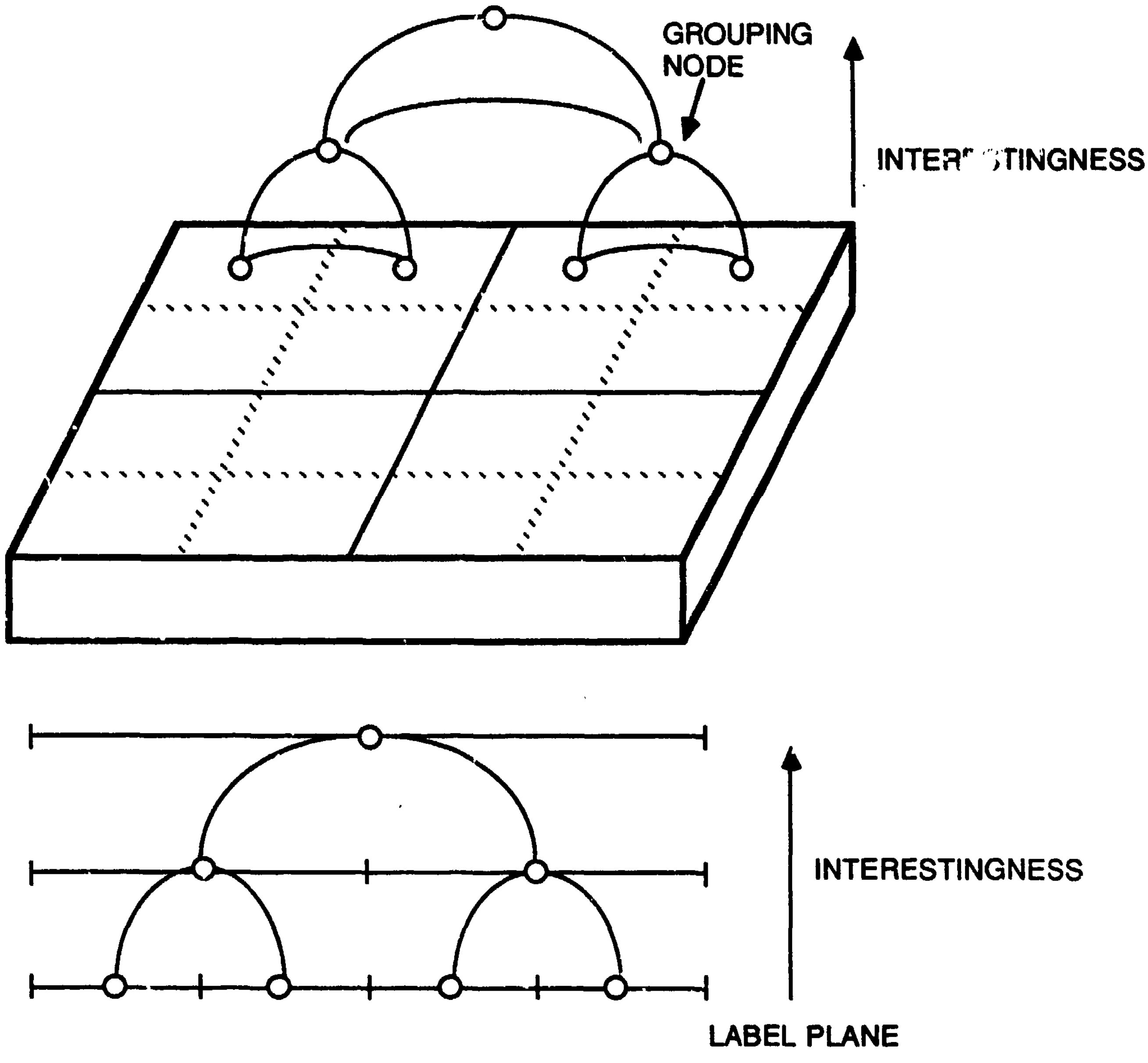


Figure 5-5: Grouping Architecture

In many systems, objects are simply treated as lists of attributes that are matched against extracted image features. Here they are treated as specifying an active control process that directs image segmentation by specifying grouping procedures to extract and organize image structures.

The process of schema instantiation creates an instance of a schema together with evidence for that schema. Evidence consists of structures in the PSDB, a priori knowledge stored in the LTM, predictions derived from location inference, and relations to already instantiated schema.

Table 5-1 shows the various slots and relationships in a generic schema. Although this data structure has a frame-like appearance, it is useful to view the schema as a semantic net structure, with slots representing nodes in the net and relationships representing arcs. Schema instantiation inference reasons from a (partially) instantiated node, follows arcs, and infers procedures to execute from the sum of its acquired information in order to obtain more evidence to further instantiate the schema.

The schema network is a generic set of data structures that indicate the a priori relationships between schemas. A key part of this network is the inheritance hierarchies that indicate which descriptions and relationships can be inherited from schema to schema. Inheritance hierarchies allow efficient matching of objects in the world against sensor evidence from progressively coarser to finer levels. As reasoning moves from coarser to finer levels of description in model-based schema instantiations, the schemas inherit descriptive bounds and add new descriptions, and also add constraints to inherited ones. For example, the system may first recognize an object as a naturally occurring ("NATURAL" in the IS-A tree) terrain patch (because of it's lack of structure and size). A "RIVER" is a type of "NATURAL" object (see Figure 4-2), that specifies linear boundary descriptions and constrains the spectral properties of the object as it should appear in the image. The two basic types of schema network inheritance hierarchies are IS-A and PART-OF, as described in Section 4.

Below is a brief explanation of each of the slots and relationships in the (preliminary) generic schema data structure. Schema type refers to the generic name of the schema in the IS-A hierarchy. Schema name is the identification of the schema instance, e.g., if the schema type is "road" then the schema name might be "highway 101". The schema instantiation structure maintains the control history of the schema recognition inference processes for this schema.

The feature description is an object-centered view of the world object represented by the schema. It includes its geometry and shape description, actual size, and contrast range and texture. Note that this is the description that matches the schema-object before looking at its structure refined into components. For example, geometric description of a "PRIMARY ROAD" schema does not

Table 5-1: Generic Schema Data Structure

- SCHEMA TYPE
- SCHEMA NAME
- SCHEMA INSTANTIATION STRUCTURE
- FEATURE DESCRIPTION
  - o SHAPE
  - o SIZE
  - o EXPECTED CONTRAST
  - o TEXTURE
- PERCEPTUAL STRUCTURE
- COMPONENTS
  - o MUST HAVE
  - o MAY HAVE
  - o VIEW DEPENDENT RELATIONSHIPS
- PART OFS
- CLASSIFICATIONS
  (POINTS UP THE IS-A HIERARCHY ONE LEVEL)
- CONTEXTUAL RELATIONSHIPS
  - o ALWAYS OCCURS WITH
  - o SOMETIMES OCCURS WITH
  - o NEVER OCCURS WITH
  - o CONFUSED WITH
  - o SIMILAR TO
- LOCATIONAL INFORMATION
  - o REGISTRATION PARAMETERS
  - o GRID AFFIXMENTS
  - o MAP LABELS
- RECOGNITION STRATEGIES

separate the actual road surface and the median strip, but gives a single enclosing area as its representation. The areal descriptions of the road surface and the median strip appear as the geometric descriptors on their schema further down the PART-OF hierarchy. Thus, inferring down the PART-OF hierarchy corresponds to increasing the resolution of the view of the object represented by the schemas.

The perceptual structure is the dynamically created PSDB query history generated by the schema instantiation as it attempts to fill in evidence matching the various schema slots and relations. The instantiator can re-use successful branches of perceptual structures to improve its recognition speed as it continues to view other instances of the same generic schema type.

Components are pointers to other schema that represent sub-parts of the schema object. They are finer resolution description of the schema, one level down on the PART-OF hierarchy. The MUST-HAVE components are assumed to be parts the represented object must have to exist, although the schema may be instantiated without observing them all. Occasionally occurring components, such as median strip on primary roads, can be stored in the MAY-HAVE slot. Spatial relationships between components as they make up the schema object are listed at this level also. PART-OF's point upward one level on the PART-OF hierarchy, indicating that this schema is a component of another schema.

Classification points upward and downward one level on the IS-A hierarchy. There may be more than one such pointer, which is to say that the IS-A hierarchy may be partially ordered.

Contextual relationships indicate spatial/temporal consonance or disconsonance between groups of schema types, omitting those which are already indicated the PART-OF and IS-A hierarchies. Schema that ALWAYS or never-occur with the given one can be used strongly for belief or disbelief in the schema instance and as focus of attention mechanisms within the instantiation process. SOMETIMES occurs with relationships that are used to store the spatial-temporal aspects of schemas relative appearance in the viewed environment.

CONFUSED-WITH and SIMILAR-TO relationships indicate schema that may be mistaken for the given one, but for different reasons. One schema may be confused with another because they share common evidence pieces, but for which there are sufficient descriptors to disambiguate. Two schema are similar if there is sufficient ambiguity in their appearances, and therefore the available perceptual evidence, that they may be indistinguishable without contextual reasoning. For example, small roads may be confused railroads from coarse shape and spectral evidence, but can often be disambiguated by other features or higher resolution data. On the other hand, roads are similar to runways because they cannot necessarily be distinguished by their intrinsic appearance, no matter how detailed or

accurate the descriptors and evidence (because they are constructed the with the identical materials, etc.). Contextual reasoning, e.g., the presence of aircraft on the runway, global curvature of the road, etc. is required.

Locational information points at the image frame(s) the schema appears in, the corresponding map identifier, and any necessary registration information.

Recognition strategies are prioritization cues for the schema instantiation processes that suggest inference chains likely to pay off to match this schema instance against sensor evidence.

The recognition strategies slot in the schema data structure prioritizes inference approaches relevant to this schema. These approaches include search for components, search for part of schema instance, search on weaker classification, relations with other schema instances, and PSDB matching.

Search for COMPONENTS and search for PART-OF are both inferences along the PART-OF hierarchy in different directions. The instantiator searches the relevant slot to see if there are components to search for or another object of which this schema is a component. If the COMPONENT or PART-OF schemas exist, they can be accessed to continue the inference. Otherwise, each causes an instantiation of the missing schema to be generated as a prediction. Instantiation control can be transferred at this point to the COMPONENT or PART-OF schema. The schema inference process maintains its thread of reasoning relevant to the schema in the schema instantiation structure slot.

5.3 DIGITAL TERRAIN DATABASE

The digital terrain database is part of LTM. It stores the data necessary for predictions of terrain features and landmark locations. The long term terrain database contains a priori map data including terrain feature representations, elevation data, and schemas representing instances of stable terrain object hypotheses extracted from the SAR data. The a priori map and elevation data (if available) is used to predict instances of terrain features and to help guide image segmentation.

In an effort to understand the issues related to the management of an evolving terrain database, this contract will investigate the availability and appropriateness of Geographic Information Systems to the autonomous extraction of features from SAR Imagery.

Eventually, the digital terrain database should be populated using information provided by data in one of the standard DMA product forms, such as Digital Terrain Elevation Data (DTED), Digital Feature Attribute Data (DFAD), or digital Tactical Terrain Attribute Data Base (TTADB) data. Once the image has

been registered to the map data, queries will be made to extract terrain features that are likely to be visible in the SAR image. This information will then be used to guide the schema instantiation processes.

As hypotheses acquire enough evidence, they are eventually moved into LTM where they in turn will be used as a priori knowledge for the next set of imagery to be exploited. In this manner, terrain databases can be built up for areas where no existing digital terrain maps were previously available. This process describes the "bootstraping" that will be necessary for exploitation of unmapped areas.

6. PROJECT STATUS

6.1 PROJECT PLAN

The goal of the Linear Feature Extraction Phase II SBIR is to develop an automated linear feature extraction system for radar imagery.

The major steps in achieving a capable linear feature extraction system are as follows:

1. Develop the appropriate working environment to register, manipulate, and process imagery
2. Develop and experiment with various segmentation and feature extraction algorithms
3. Determine significant terrain object feature properties and construct representative object models
4. Experiment and evaluate model to image feature matching schemes
5. Develop an approach for managing the competing and conflicting hypothesis matches
6. Develop feature finders/predictors to support or contradict an expected terrain feature's existence.
7. Implement a display interface to support the above processing steps.

This project is divided into three parts.

Base Contract - (6 months) Undertake and complete the design of an automated linear feature extraction system for SAR imagery.

Option I - (9 months) Undertake and complete the development of all necessary software for the core system components of such a system. Work will also begin for recognition technique development and the system development. (This option overlaps the previous phase by three months.)

Option II - (12 months) Complete the work on the recognition technique development and the system development work began in the previous effort.

6.2 REVIEW OF PROGRESS

The work performed by ADS under the Base Contract has addressed three different problem areas.

The primary work performed under this contract was the continuation of the design produced in the Phase I SBIR effort. The results of that design are described in Sections 3, 4, and 5 of this document.

The second major area in which ADS pursued the project goals was the development and the design of a software environment in which to perform experiments and begin to build the eventual prototype system. The basic framework of this software was delivered to ETL in May. The delivery emphasized neighborhood and display operations. The software also contained the necessary software "hooks" for future expansion into the other system components.

Finally, the last area of work undertaken as part of the Basic Contract was the continued experimentation with the government provided radar imagery. Experimentation included algorithm surveys, hand processing sample imagery, and actual algorithm implementation. This work and ADS's general understanding of machine vision, has been continually supporting the design and development of the components of a model based vision system for linear feature extraction. The work described above corresponds to significant progress in Steps 1, 2 and 7 and has established the infrastructure for continuing work on the other steps of the Project Plan.

Once the proper environment is established, this system for determining and extracting terrain features can be extensively tested. These experiments will further establish the role of autonomous feature extraction from SAR imagery and, indeed, the importance of SAR imagery to map generation.

7. REFERENCES

[Hanson et.al. - 78] - A.R. Hanson and E.M. Riseman, "VISIONS: A Computer System for Interpreting Scenes", In Computer Vision Systems, Academic Press, 1978.

[Lawton et. al. - 85] - D.T. Lawton, J. Glicksman, G.D. Conner, R.J. Drasovich, "Linear Feature Extraction from Radar Imagery", Final Technical Report, TR-3075-01, Contract no. DACW72-84-C-0014, Advanced Decision Systems, Mountain View, California, August 1985.

[Marr - 82] - D. Marr, "Vision", W.H. Freeman, San Francisco, 1982.

[Martelli - 76] - A. Martelli, "An application of heuristic search methods to edge and contour detection", Commun. ACM, Vol. 19, No. 2, February 1976, pp. 73-83.

[Ohta - 80] - Y. Ohta, "A Region-Oriented Image-Analysis System by Computer", Ph.D. Thesis, Kyoto University, Department of Information Science, Kyoto, Japan, 1980.